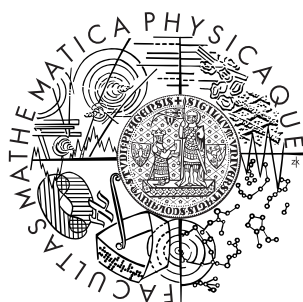


Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



Bc. Viktor Mašíček

Integrace Sociálních sítí

Katedra softwarového inženýrství

Vedoucí diplomové práce: Mgr. Jaroslav Tykal

Studijní program: Informatika, ISS

2009

Děkuji Mgr. Jaroslavu Tykalovi za vstřícnost a čas, který mi ochotně věnoval, i za rady a náměty, které mi byly cenným vodítkem při realizaci této práce.

Prohlašuji, že jsem svou diplomovou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

V Praze dne 7.12.2009

Bc. Viktor Mašíček

Obsah

1	Úvod	8
1.1	Cíl práce	8
1.2	Proč integrovat	9
1.3	Co je obsahem	9
2	Seznámení s pojmy	11
2.1	Co to je sociální síť	11
2.2	Historie	11
2.3	Typická struktura	12
2.3.1	Profil	12
2.3.2	Kontakty	13
2.4	Sociální graf	13
2.5	Shrnutí	14
3	Rozbor existujících sociálních sítí	15
3.1	Seznam sítí a jejich popis	15
3.1.1	aNobii	15
3.1.2	CouchSurfing	16
3.1.3	Facebook	16
3.1.4	Friendster	16
3.1.5	Last.fm	16
3.1.6	Lidé	16
3.1.7	LinkedIn	17
3.1.8	MySpace	17
3.1.9	PatientsLikeMe	17
3.1.10	ResearchGate	17
3.2	Vlastnosti sítí a jejich popis	17
3.2.1	Skupiny uživatelů (skupiny)	18

3.2.2	Psaní zpráv mezi uživateli (zprávy)	18
3.2.3	Chatování mezi uživateli (chat)	18
3.2.4	Diskusní fóra (fórum)	18
3.2.5	Nahrávání fotografií (fotky)	18
3.2.6	Nahrávání videí (video)	18
3.2.7	Import kontaktů z emailu (import)	19
3.2.8	Přehled vlastností zmíněných sítí	19
3.3	Typy sítí	19
3.3.1	Zaměření	20
3.3.2	Možnost založení účtu	20
3.4	Shrnutí	21
4	Snahy o integraci	22
4.1	Friend of a Friend	22
4.2	Xhtml Friends Network	23
4.3	Social Graph API	23
4.4	OpenSocial	23
4.5	OpenId	24
4.6	Shrnutí	24
5	Co chceme integrovat	25
5.1	Profil	25
5.2	Seznam kontaktů	26
5.3	Vlastnosti	26
5.4	Shrnutí	27
6	Způsob integrace	28
6.1	Závislost vs. nezávislost	28
6.2	Konektory	29
6.3	Hlavní sociální síť	30
6.4	Shrnutí	30
7	Návrh integrace	32
7.1	Procesy v Hlavní síti	32
7.1.1	Registrace uživatele	32
7.1.2	Synchronizace profilu	33
7.1.3	Synchronizace seznamu kontaktů	33
7.1.4	Návrhy skupin	34

7.1.5	Pravidelná kontrola uživatelů	34
7.2	Identifikace problémů a návrh řešení	34
7.2.1	Ztotožnění atributů profilů ve vzdálených sítích	34
7.2.2	Uživatel chce rozdílné profily	35
7.2.3	Jednoznačná identifikace uživatele	35
7.2.4	Uživatel chce rozdílné seznamy kontaktů	35
7.2.5	Automaticky založený uživatel	36
7.2.6	Detekce rozdílů v seznamech kontaktů	36
7.2.7	Návrh nových skupin	36
7.3	Návrh datového modelu	37
7.3.1	Souhrn požadavků	37
7.3.2	Uživatelé, sítě a členství	38
7.3.3	Neregistrovaní uživatelé	38
7.3.4	Nastavení	39
7.3.5	Atributy profilů	40
7.3.6	Profily	41
7.3.7	Identifikátory	42
7.3.8	Seznamy kontaktů	42
7.3.9	Skupiny	42
7.3.10	Celý model	44
7.4	Shrnutí	45
8	Závěr	47
9	Budoucí práce	50
	Literatura	51

Název práce: Integrace Sociálních sítí

Autor: Bc. Viktor Mašíček

Katedra: Katedra softwarového inženýrství

Vedoucí diplomové práce: Mgr. Jaroslav Tykal

E-mail vedoucího: Jaroslav.Tykal@mff.cuni.cz

Abstrakt: Sociální sítě jsou současným fenoménem a jejich integrace začíná získávat na důležitosti. Základem této integrace je spárování shodných informací v jednotlivých sociálních sítích a detekce nesrovnalostí mezi nimi. Nejdůležitější data, která je potřeba integrovat, jsou profily uživatelů a seznamy jejich kontaktů.

V naší práci jsme navrhli integraci jak profilů a seznamů kontaktů, tak skupin uživatelů, které je možné v sociálních sítích vytvářet. Součástí práce není implementace. Návrh však počítá s Hlavní sociální sítí, která by webovým rozhraním zprostředkovala integraci koncovým uživatelům. Z pohledu uživatelů je přínosné zobrazení dat ze všech sociálních sítí na jenom místě a automatizovaná detekce rozdílů.

Data uživatelů by pak mohla být využita pro obchodní účely, samozřejmě v rámci zákonných omezení. Z tohoto pohledu je největším přínosem Hlavní sociální sítě sociální graf. Ten by se skládal ze sociálních grafů jednotlivých sociálních sítí. Navíc by v něm byly informace o členství uživatelů v jednotlivých sociálních sítích.

Součástí návrhu jsou hlavní procesy Hlavní sociální sítě, její datový model a způsob získávání informací z jednotlivých sítí. Při návrhu modelu jsme se částečně inspirovali existujícími projekty. Při získávání informací ze sociálních sítí by se dokonce některých projektů dalo využít v plné míře.

Klíčová slova: sociální síť, sociální graf, sémantický web

Title: Social networks integration

Author: Bc. Viktor Mašiček

Department: Department of Software Engineering

Supervisor: Mgr. Jaroslav Tykal

Supervisor's e-mail address: Jaroslav.Tykal@mff.cuni.cz

Abstract: Social networks are the current phenomenon, and their integration begins to gain importance. Basic idea is to pair identical pieces of information stored in various social networks and to detect inconsistencies between them. The most important data that need to be integrated are user profiles and lists of contacts.

In our work, we propose the integration of both the profiles and lists of contacts, and even groups of users that could be created in social networks. Implementation is not part of the work. However the proposal suggests the creation of Main social network in form of a web application, that brokers the integration to end users. From the perspective of users it is beneficial to display data from all social networks in one place and automate detection of differences.

User's data could be used for commercial purposes, of course, within legal limits. From this perspective, the biggest contribution of Main social network is a social graph. It consists of various social graphs of social networks. Additionally it would contain information about membership of users in social networks.

The proposal consists of three main parts: main processes of the Main social network, its data model and means of gathering information from different networks. The design of data model was partially inspired by existing projects. Some projects have already addressed the gathering of data, so they could be fully used.

Keywords: social network, social graph, semantic web

Kapitola 1

Úvod

Tématem této práce jsou sociální sítě. Jedná se o určité komunity lidí, kteří mají nějaký společný zájem. Díky sociálním sítím je možné sdílet data nebo informace, poznávat nové přátele nebo navazovat obchodní vztahy. Těchto komunit v poslední době velmi přibývá. Jelikož se zaměřují na různá témata, mnoho uživatelů má vytvořeno řadu profilů v několika sociálních sítích. Některé informace o uživateli se pak nachází redundantně ve více sítích a je mnohdy těžké udržet informace na více místech aktuální.

1.1 Cíl práce

Cílem práce je zmapovat existující sítě a zjistit, jaké mají společné rysy, případně zda je nelze rozdělit do různých kategorií nebo typů.

Na základě získaných informací je úkolem přednést návrh na integraci, pokud možno všech typů sítí, nezávisle na jejich zaměření, obsahu nebo čemkoli jiném. Mělo by se tedy jednat o *Hlavní sociální síť*, která bude zastřešovat všechny ostatní. Tato Hlavní síť by měla navzájem mapovat logicky shodné informace uložené redundantně v ostatních sítích.

Přínosem Hlavní sítě by mělo být vytvoření mnohem komplexnějších profilů jednotlivých uživatelů. Pokud by se navíc vytvořil patřičný interface, bylo by pro uživatele spravování účtů ve více sítích mnohem pohodlnější.

Samotná technická realizace Hlavní sítě však již překračuje rámec této práce. Cílem je pouze analýza a návrh Hlavní sítě.

1.2 Proč integrovat

Otázkou je, proč bychom měli zájem na integraci všech sociálních sítí do jedné společné. Na důvody můžeme nahlížet jak z pohledu uživatelů, tak z pohledu provozovatelů Hlavní sítě.

Z pohledu uživatelů je jeden ze základních důvodů přehlednost a jednoduchost. V různých sítích mají uživatelé uloženy různé informace. V Hlavní síti budou mít všechny pod kontrolou na jednom místě. Uživatel pak není nucen se přihlašovat do všech sítí, ve kterých je registrován.

Dalším přínosem je kontrola nesrovnalostí. V různých sítích mohou být uloženy stejné informace. Pokud dojde ke změně, mohlo by se stát, že uživatel provede opravu jen v jedné síti. Hlavní síť tuto nesrovnalost detekuje a uživatel může provést nápravu.

Naopak z pohledu provozovatelů je zřejmý ekonomický přínos. Uživatelé v jednotlivých sítích uvádějí informace o sobě, svých přátelích atd. Provozovatelé Hlavní sítě pak mají přehled o preferencích jednotlivých uživatelů. Navíc budou mít preference rozděleny na základě zdrojů informací, tedy z jaké sítě jsou jednotlivé informace získány. V Hlavní síti budou také informace o mnohem větším počtu uživatelů. V souhrnu tak provozovatelé budou mít možnost vytvořit mnohem komplexnější obrázek o preferencích uživatelů. Díky tomu pak mohou nabízet uživatelům lépe cílené reklamní nabídky. To je samozřejmě z obchodního pohledu přínosné jak pro provozovatele, tak i pro uživatele.

1.3 Co je obsahem

V kapitole 1 je popsáno, co je cílem této práce a jaké jsou důvody pro integraci. V závěru je stručný popis jednotlivých kapitol celé práce.

V kapitole 2 se seznámíme s pojmem *Sociální síť*. Poté je zde něco o historii, popis typické struktury sociálních sítí a v závěru kapitoly je vysvětlen pojem *Sociální graf*.

V kapitole 3 vyjmenujeme některé existující sociální sítě a stručně si je popíšeme. Poté vyjmenujeme a popíšeme některé z doplňujících vlastností sociálních sítí. Následně je zobrazeno, u kterých vyjmenovaných sítí se které vlastnosti vyskytují. V závěru rozebereme, jak lze dělit sítě podle kritérií.

V kapitole 4 si popíšeme některé projekty, které se již integrací zabývají.

V kapitole 5 je uveden výčet toho, co je pro nás zajímavé integrovat.

V kapitole 6 rozebereme jaký směr zvolíme pro naši integraci. Vzhledem k výběru který zvolíme, jsou zde také podkapitoly věnované pojmům *Konektor* a *Hlavní síť*.

V kapitole 7 identifikujeme, jaké jsou hlavní procesy probíhající v Hlavní síti. Následně identifikujeme problémy a navrhneme jejich řešení. V závěru navrhneme datový model na základě požadavků vzniklých v předchozím textu.

V kapitole 8 je shrnutí celé práce.

V kapitole 9 popíšeme několik možností, jak dále pokračovat v námi nastíněné i integraci.

Kapitola 2

Seznámení s pojmy

Před návrhem Hlavní sítě si nejprve řekneme, co to vlastně sociální sítě jsou, kde je nalezneme a jak vznikly. Pak si stručně přiblížíme charakteristické znaky sociálních sítí. V závěru kapitoly se seznámíme s pojmem Sociální graf.

2.1 Co to je sociální síť

Sociální sítě se chápou jako společenství lidí, mezi kterými jsou určité vztahy. Může se jednat o vztahy přátelské, obchodní, pracovní atd. Internet a Web, tak jak je známe dnes, umožňují vytvářet sociální sítě mezi lidmi, kteří se dokonce ani fyzicky nesetkali. Tito lidé pak mohou navzájem sdílet informace nejrůznějších druhů.

V našem chápání je tedy sociální síť nějaká internetová služba, zpravidla zprostředkovaná webovým rozhraním, pomocí které je možné vytvářet vazby mezi lidmi a sdílet informace.

2.2 Historie

Pojem sociální síť je používán až v posledních letech, ale zárodky tohoto fenoménu je možné nalézt mnohem dříve. Již síť ARPANET [1], ze které později vznikl Internet, by bylo možno považovat za první sociální síť. Na začátku spojoval několik amerických univerzit a později se přidávaly další (nejen americké). Společným zájmem těchto univerzit byl právě samotný ARPANET a

výzkum sítí. Již od počátku byly do ARPANETU zapojovány také další projekty, které mají k sociálním sítím hodně blízko. Jako například Usenet [2], který byl předchůdcem dnešních diskusních fór.

V devadesátých letech začaly vznikat nové sociální sítě. První, která obsahovala všechny možnosti předchozích, byla SixDegrees [3]. Bylo možné zakládat nové členy a editovat jejich profily, vytvářet přátelské vazby a vyhledávat nové přátele dle charakteristik v jejich profilech. Toto je již model, který se velice podobá sociálním sítím tak, jak je známe dnes.

Po roce 2000 vznikly velmi populární sítě Friendster [4] a MySpace [5]. Dle [6] byla v roce 2005 MySpace největší ze všech sociálních sítí a údajně měla větší návštěvnost než Google. V současné době je velké množství sociálních sítí a stále přibývají další. Zaměřují se na nejrůznější oblasti jako hudba, knihy, obchod, zaměstnání atd. Velmi obsáhlý a hlavně kategorizovaný přehled sociálních sítí se nachází na [7]. Naopak na [8] se nachází přehled těch nejznámějších sítí včetně přibližného počtu uživatelů a stručného popisu.

2.3 Typická struktura

Převážná většina sociálních sítí má velmi podobné rysy. Základem je, že uživatel může vyplnit svůj profil a vytvářet si seznam kontaktů (přátel, obchodních partnerů, kolegů atd.). Mohou ještě existovat mnohé další funkcionality jako sdílení fotek nebo odkazů, chatování, vytváření skupin uživatelů atd. Ty se však nemusí vyskytovat ve všech sociálních sítích a nejsou nezbytnou součástí pro jejich fungování. Záleží na účelu sítě a komfortu, který chce autor uživatelům nabídnout.

2.3.1 Profil

Nejzákladnější vlastností sociálních sítí je vytvoření vlastního profilu. Jeho obsah se může lišit síť od sítě. V některých je možné zadat jen základní údaje, v jiných například i jaké školy uživatel studoval, kde pracuje a pracoval, jaké má koníčky atd. Povinné ale bývají jen základní údaje, v některých případech jen email.

V sítích, ve kterých je možné o sobě zadávat větší množství informací do profilů, můžeme většinou také určit, jaké informace komu zveřejnit. Například jméno zveřejnit všem (i neregistrovaným osobám), email jen registrovaným uživatelům a svoje telefonní číslo jen uživatelům ze svého seznamu kontaktů.

2.3.2 Kontakty

Ve všech sítích je možné hledat ostatní uživatele a přidávat si je do seznamu svých kontaktů, mnohdy označovaných jako přátelé. Tam, kde to nelze, se nejedná o sociální síť.

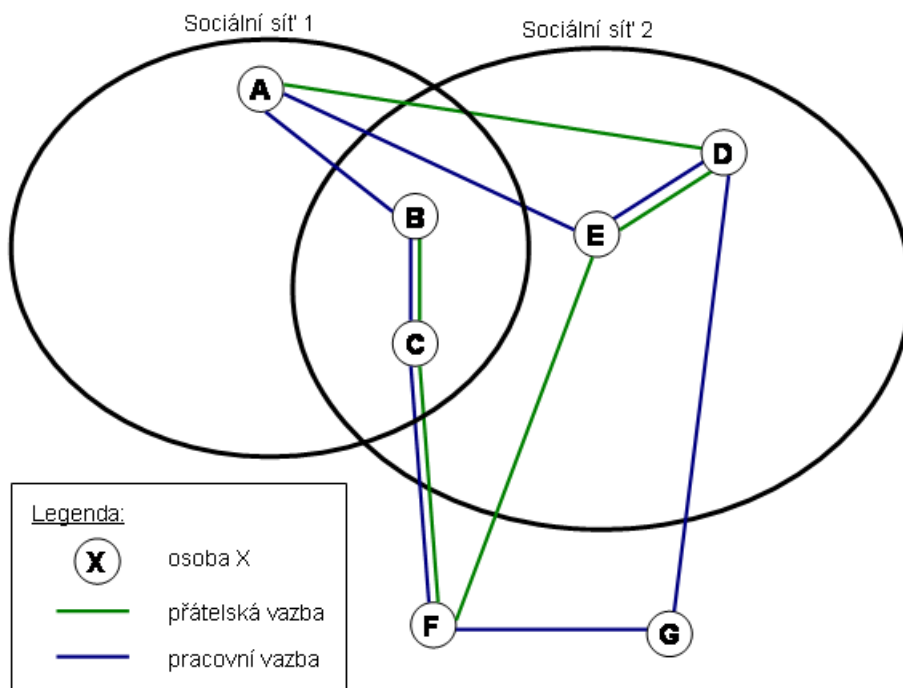
Obvykle platí, že jiného uživatele si nelze přidat mezi své kontakty bez jeho souhlasu. Přidaný uživatel musí potvrdit, že chce být ve vašem seznamu kontaktů. V ojedinělých případech to však není nutné. Například na ČSFD [9] si lze ostatní uživatele bez jejich souhlasu přidat do svého seznamu oblíbených uživatelů. Na druhou stranu primární funkcí zmíněného serveru není vytváření kontaktů mezi lidmi, takže by se dalo polemizovat, zda se vlastně jedná o sociální síť.

2.4 Sociální graf

Mezi lidmi existují ve skutečném světě sociální vazby. Někteří lidé mají mezi sebou pracovní vztahy, někteří kamarádské atd. V sociálních sítích máme tyto vazby také. Velmi často vznikají na základě vazeb ze skutečného světa. Někdy však mohou vznikat až v rámci sociálních sítí. Tyto vazby se automaticky stávají součástí skutečného světa, i když se jednotliví lidé nemuseli nikdy fyzicky potkat.

Ve skutečném světě představuje sociální graf vazby mezi jednotlivými lidmi. V rámci jednotlivých sítí sociální graf představuje vazby mezi jednotlivými uživateli (kap. 2.3.2). Je důležité si uvědomit, že sociální graf z libovolné sítě je podgrafem sociálního grafu ze skutečného světa. A také, že sociální grafy z jednotlivých sítí se mohou překrývat, a velice často tomu tak je. Pro lepší představu je na obrázku 2.1 zobrazen příklad sociálního grafu reálného světa zasahujícího do dvou sociálních sítí.

Snaha o integraci je právě snaha o vzájemné využití informací ze sociálních grafů různých sítí. Problémem je, že právě sociální graf je pro provozovatele sociálních sítí jednou z nejcennějších informací, převážně z marketingového hlediska. A proto v řadě sítí nebývají sociální grafy přímo dostupné.



Obrázek 2.1: Na obrázku vidíme jednotlivé osoby. Jsou zde také znázorněny jejich vazby. Jak je vidět, mezi osobami může být více vazeb. Zde jsou znázorněny přátelské vazby (zelená) a pracovní vazby (modrá). Osoby **A**, **B** a **C** společně s jejich vzájemnými vazbami tvoří sociální graf *Sociální síť 1*. Osoby **B**, **C**, **D** a **E** společně s jejich vzájemnými vazbami tvoří sociální graf *Sociální síť 2*. Celý obrázek pak znázorňuje sociální graf skutečného světa. Pro jasné pochopení uvedeme, že například vazby mezi **A** a **D** nebo **F** a **G** patří pouze do sociálního grafu reálného světa. To proto, že koncové osoby těchto vazeb nejsou členové stejných sociálních sítí. Kdybychom se však ptali na sociální graf *Sociální síť 1* a *Sociální síť 2* dohromady, vazba mezi **A** a **D** by sem patřila.

2.5 Shrnutí

V této kapitole jsme si vysvětlili některé základní pojmy. Zaprvé, že za sociální síť považujeme jakoukoli webovou službu, ve které je možné vyplňovat svůj profil a vytvářet vazby s ostatními uživateli. Zadruhé, že grafu, který je tvořen vazbami mezi všemi uživateli v nějaké sociální síti, říkáme sociální graf.

Kapitola 3

Rozbor existujících sociálních sítí

Sociálních sítí existuje velké množství. Nejprve si některé nejznámější představíme. Také se zmíníme o pár ne příliš známých, ale poměrně zajímavých, a stručně si popíšeme, co je jejich hlavním cílem.

Velmi často jsou sociální sítě rozšířeny o různé vlastnosti. Tyto vlastnosti nejsou nezbytnou podmínkou pro fungování sítě, nicméně zlepšují uživatelský komfort. V dalším textu si proto zmapujeme vlastnosti zmíněných sítí a uvedeme jejich přehled.

V závěru kapitoly si pro lepší orientaci a vhodnější návrh integrace rozebereme, podle kterých kritérií lze sociální sítě rozlišovat.

3.1 Seznam sítí a jejich popis

V této kapitole uvedeme výčet některých známých a zajímavých sociálních sítí. Pro přehlednost budou jednotlivé sítě uvedeny v abecedním pořadí.

3.1.1 aNobii

Síť aNobii [10] se zaměřuje na knihy. Máme možnost si vytvářet seznam knih ve své knihovničce a zjišťovat oblíbenost knih. Velmi zajímavou funkcí je lokalizace knihkupectví a knihoven na mapě a jejich vyhledávání podle nejrozličnějších kritérií.

3.1.2 CouchSurfing

CouchSurfing [11] je síť zaměřená na cestování. Její členové si navzájem nabízejí ubytování po celém světě. Velmi často bývá součástí ubytování také provedení po památkách a zajímavostech okolí. Každý člen ve svém profilu napíše, jaké má jazykové znalosti a jaké ubytování nabízí. Ostatní pak mohou podle svých požadavků vyhledat nejlepší nabídky. Po využití ubytování je možné napsat zhodnocení pro další uživatele.

3.1.3 Facebook

Facebook [12] je v dnešní době jedna z největších sociálních sítí na světě co do počtu uživatelů, podle [8] je druhá největší. V poslední době je velmi oblíbená i v České republice. Hlavním cílem je udržovat aktuální kontakty na přátele. Facebook také podporuje speciální aplikace, což umožňuje zlepšovat jeho funkcionalitu.

3.1.4 Friendster

Friendster [4] patří, co do počtu uživatelů, k větším sociálním sítím. Jejími uživateli jsou převážně obyvatelé států ze sdružení ASEAN [13]. Síť nemá žádné konkrétní zaměření. Jejím cílem je sdružovat přátele a umožňovat výměnu kontaktů, sdílení fotografií atd.

3.1.5 Last.fm

Last.fm [14] je hudebně zaměřená síť. Je možné na ní poslouchat rádia, do kterých se sestavují písničky podle vašeho vkusu, nebo například vyhledávat hudební akce v okolí vašeho bydliště. Také je možné vyhledávat uživatele s podobným hudebním vkusem.

3.1.6 Lidé

Lidé [15] je jedna z největších sociálních sítí v České republice. Zaměřuje se převážně na diskuze, chaty, seznamky a služby podobného charakteru.

3.1.7 LinkedIn

LinkedIn [16] je jedna z největších sociálních sítí zaměřených na kariéru a byznys. Je určena k hledání pracovních příležitostí a nových pracovníků, ale také samostatných obchodních příležitostí.

3.1.8 MySpace

MySpace [5] je největší sítí, co do počtu uživatelů (podle [8]). Sloganem sítě je “Místo pro přátele”, což naznačuje její hlavní funkci. Hledání přátel, udržování kontaktů a zábava. Tomu také odpovídá podpora velkého množství vlastností a speciálních aplikací.

3.1.9 PatientsLikeMe

PatientsLikeMe [17] je síť určená pro pacienty trpící některou z vážnějších dlouhodobých nemocí (např. HIV, epilepsie, . . .), a také pro lékaře a výzkumníky, zabývající se těmito nemocemi. Je možné navazovat kontakty s dalšími pacienty nebo s lékaři a vyhledávat výsledky výzkumů v dané oblasti.

3.1.10 ResearchGate

ResearchGate [18] je síť určená pro vědecké pracovníky. Mají možnost zde zveřejňovat své výzkumy a články. Velký přínos spočívá v hledání ostatních vědců pro navázání případné spolupráce.

3.2 Vlastnosti sítí a jejich popis

V této kapitole rozebereme vlastnosti, které se u sociálních sítí vyskytují. V závěru kapitoly bude přehled vlastností vyjmenovaných sociálních sítí (kap. 3.1). Na vlastnosti se budeme odkazovat jednoslovnými hesly. Tato hesla budou uvedena v závorce v názvu kapitoly každé vlastnosti.

Nebudeme se rozepisovat o každé jednotlivé vlastnosti, protože téměř každá síť má například rozšiřující možnosti profilu nebo seznamu kontaktů. My se zaměříme na ty nejzajímavější a nejvíce rozšířené.

3.2.1 Skupiny uživatelů (skupiny)

Většina sítí má velké množství uživatelů. Tento fakt znesnadňuje nalezení nových kontaktů, které by měly podobné zájmy. Některé sítě však podporují vytváření skupin. Jednotlivé skupiny jsou vytvářeny za účelem sdružit uživatele s podobnými zájmy. Uživatelé pak mají možnost se do skupin přihlašovat, což hledání nových kontaktů velmi usnadňuje.

3.2.2 Psaní zpráv mezi uživateli (zprávy)

I když prakticky všechny sítě umožňují zadávat do profilu uživatele jeho email, ne každý této možnosti využije. Pokud tedy chcete napsat někomu zprávu, je možné to provést přímo v rámci rozhraní sítě. Velmi často je možné zaslat zprávu i více uživatelům.

3.2.3 Chatování mezi uživateli (chat)

Řada sítí je rozšířena o možnost chatování mezi uživateli. Někdy pouze mezi jednotlivými uživateli, někdy je možné vytvářet chatovací místnosti.

3.2.4 Diskusní fóra (fórum)

Dalším komunikačním kanálem, který bývá podporován, je diskusní fórum. Velmi často bývá spojen s uživatelskými skupinami (kap. 3.2.1), ale není to podmínkou.

3.2.5 Nahrávání fotografií (fotky)

Velmi populární je možnost nahrávání fotografií, někdy rozšířená o vytváření fotoalb. Tyto fotky je pak možné zpřístupnit ostatním uživatelům, zpravidla těm ve svém seznamu kontaktů.

3.2.6 Nahrávání videí (videa)

Některé sítě umožňují také nahrávání videí. Není to však tak rozšířená vlastnost, jelikož je náročná na kapacitu disků na serveru.

3.2.7 Import kontaktů z emailu (import)

Velmi užitečná funkce je vyhledání kontaktů na základě seznamu adres z emailové schránky. V rozhraní sociální sítě zadáte login a heslo a zobrazí se vám seznam těch, které máte mezi kontakty v emailové schránce, a zároveň jsou členy příslušné sítě. Zpravidla jsou podporovány emailové servery jako Gmail, Yahoo!, Hotmail atd.

3.2.8 Přehled vlastností zmíněných sítí

V tabulce 3.1 je přehled sítí z kapitoly 3.1 a jejich vlastností. Existence vlastnosti v rámci dané sítě je vyznačena znakem '*' v příslušné kolonce.

U všech zmiňovaných sítí předpokládáme, že mají základní vlastnosti, tedy možnost administrace profilu a možnost vytváření seznamu kontaktů z ostatních uživatelů (kap. 2.3).

	SKUPINY	ZPRÁVY	CHAT	FÓRUM	FOTKY	VIDEO	IMPORT
aNobii	*		*	*			
CouchSurfing	*	*	*		*		
Facebook	*	*	*		*	*	*
Friendster	*	*	*		*	*	*
Last.fm	*	*					
Lidé	*	*	*	*	*		
LinkedIn	*	*					*
MySpace	*	*	*	*	*	*	*
PatientsLikeMe	*	*		*			
ResearchGate	*	*					*

Tabulka 3.1: Přehled vlastností sociálních sítí z kapitoly 3.1. Pokud má síť danou vlastnost, je tento fakt vyznačen znakem '*' v příslušné kolonce.

3.3 Typy sítí

Kritérií, podle kterých lze sociální sítě dělit, je celá řada. My se však zaměříme na dvě poměrně zásadní. Těmi jsou rozdělení na základě *zaměření* a na základě *možnosti založení účtu*. Uvidíme, že mezi oběma existuje určitý vztah.

3.3.1 Zaměření

Sociální sítě lze podle zaměření rozdělit do dvou hlavních skupin. Jedna skupina jsou sociální sítě bez konkrétního zaměření. Většinou jsou určeny pro udržování kontaktů mezi přáteli, sdílení fotek, odkazů, chatování atd. Těmto sítím budeme říkat *obecné*. Sociální sítě z druhé skupiny mají nějaké konkrétní zaměření. Těm budeme říkat *specializované*.

Specializované sociální sítě lze ještě dále dělit podle konkrétního zaměření. Existují sítě zaměřené na medicínu, cestování, kariéru a mnoho dalších činností.

Přehled zaměření sociálních sítí z kapitoly 3.1 je v tabulce 3.2.

	ZAMĚŘENÍ
aNobii	specializovaná - knihy
CouchSurfing	specializovaná - cestování
Facebook	obecná
Friendster	obecná
Last.fm	specializovaná - hudba
Lidé	obecná
LinkedIn	specializovaná - kariéra
MySpace	obecná
PatientsLikeMe	specializovaná - medicína
ResearchGate	specializovaná - věda

Tabulka 3.2: Přehled zaměření sociálních sítí z kapitoly 3.1.

3.3.2 Možnost založení účtu

Ne do všech sociálních sítí je možné se volně registrovat. Sítí, do kterých je možné se registrovat bez povolení administrátora, věku nad osmnáct let, přístupu k emailu s koncovkou *.edu* nebo bez nějakého jiného omezení, nazýváme *otevřené*. Sítě, do kterých není možné se registrovat bez omezení, nazýváme *uzavřené*. Otevřené sociální sítě mají samozřejmě zpravidla větší množství uživatelů než uzavřené.

Uzavřené bývají velmi často zároveň specializované (kap. 3.3.1), což je vcelku logické, protože u obecných sítí zpravidla není nutné omezovat regis-

traci uživatelů. Příkladem uzavřené sociální sítě může být síť SoberCircle [19], pomáhající alkoholikům a drogově závislým.

Naopak otevřené sociální sítě bývají jak obecné tak specializované. Například Facebook (kap. 3.1.3) je otevřená obecná síť, zatímco aNobii (kap. 3.1.1) je otevřená specializovaná síť.

3.4 Shrnutí

Jak jsme zjistili, podpora skupin uživatelů a psaní zpráv mezi uživateli je u většiny sociálních sítí, které jsme diskutovali, téměř samozřejmostí. Velmi často se také vyskytuje možnost nahrávání fotek (případně videí) a chatování.

Velmi užitečná vlastnost je import kontaktů z emailu. V ojedinělých případech by tato vlastnost byla nadbytečná (například u PatientsLikeMe, kap. 3.1.9), ale u většiny sítí přináší velké výhody, hlavně po vytvoření nového účtu.

Na závěr jsme zjistili, že sociální sítě lze rozdělit na základě dvou důležitých kritérií. Prvním z nich je *zaměření*, které dělí sítě na *obecné* a *specializované*. Druhým je *možnost založení účtu*, které dělí sítě na *otevřené* a *uzavřené*. Také jsme zjistili, že uzavřené sociální sítě bývají zpravidla specializované.

Zajímavý poznatek je, že vlastnosti sítí se jen málo odvíjejí od zaměření sítě. To se nám při integraci velmi hodí, protože nemusíme zaměření sítí zohledňovat.

Kapitola 4

Snahy o integraci

Tato práce zdaleka není první, zabývající se integrací sociálních sítí. Existuje několik aktivních projektů, věnujících se této problematice. Další podkapitoly tyto projekty popisují. Je možné se tak nejen poučit, ale také vyvarovat některých chyb.

4.1 Friend of a Friend

Projekt Friend of a Friend [20] (dále FOAF) se zabývá strojovým popisem osob, jejich aktivit a vztahů mezi nimi. Jde o formát XML řídicí se specifikací RDF¹ [21] a OWL² [22]. Řadí se tedy mezi projekty v rámci Sémantického webu [23].

Každá osoba je identifikována nějakým jednoznačným identifikátorem. Tím může být email, Jabber ID, ICQ ID, domácí webová stránka atd. Tyto identifikátory se pak využívají k vytváření vztahů mezi jednotlivými osobami nezávisle na jakékoli sociální síti. Jednotlivé sítě pak tento protokol mohou či nemusejí využívat.

Velmi praktické je v tomto projektu použití XML. Díky tomu je možné definovat profily uživatelů a jejich vztahy obecně a nezávisle na jednotlivých sítích. Nevýhodou je, že je závislý na implementaci v rámci každé sítě.

¹RDF je rodina specifikací, používaná pro modelování informací. Základní myšlenkou je trojice *podmět-vlastnost-předmět*, která k *podmětu* přiřadí *předmět* ve vztahu určeném *vlastností*. Například “auto má vysokou cenu”, kde podmět je “auto”, vlastnost “vysokou” a předmět je “cenu”.

²OWL je sémantický značkovací jazyk pro publikaci a sdílení ontologií Webu. Byl vytvořen jako slovníkové rozšíření RDF.

4.2 Xhtml Friends Network

Projekt Xhtml Friends Network [24] (dále XFN) je způsob jak reprezentovat vztahy mezi uživateli Internetu pomocí drobného rozšíření již existujícího HTML formátu. Jde pouze o přidání atributu *rel* do HTML tagu *a*. Jedná se tedy o takzvaný *mikroformát*³ a je řazen mezi projekty v rámci Sémantického webu [23].

Výhodou je jeho jednoduchost. Naopak nevýhodou je, že každý uživatel si musí pracně udržovat atributy *rel* ve všech odkazech na své kontakty na všech svých stránkách. Navíc v některých službách ani nemáme možnost atribut *rel* editovat. Jsme tedy závislí na jeho podpoře v rámci jednotlivých služeb.

4.3 Social Graph API

Projekt Social Graph API [26], od společnosti Google, je jednoduché API určené pro vývojáře. Snaží se zpřístupnit sociální vazby mezi uživateli Internetu nezávisle na jednotlivých sociálních sítích, diskusních fórech atd. Snaží se tedy o vytvoření globálního sociálního grafu (kap. 2.4). Pro vytváření vazeb a jejich detekci jsou využívány projekty FOAF (kap. 4.1) a XFN (kap. 4.2).

Zajímavou funkcionalitou je *kanonizace URL*⁴. Ta usnadňuje identifikaci jednotlivých uživatelů a jejich vzájemné propojení. Na [27] je možné stáhnout knihovny pro kanonizaci URL v různých programovacích jazycích.

Projekt si samozřejmě také nese výhody a nevýhody projektů FOAF a XFN.

4.4 OpenSocial

Projekt OpenSocial [28], od společnosti Google, usnadňuje vývojářům vytvářet nové aplikace do nejrůznějších sociálních sítí. Jde o API pro vytváření nových gadgetů⁵, postavené na klasickém HTML a SQL. Nezbytné je,

³Citace z [25]: “Mikroformát je způsob, jak do webových stránek ukládat strojově čitelné informace pomocí stávajícího HTML nebo XHTML”

⁴Kanonizace URL je převedení různých zápisů URL, směřujících na stejnou stránku, na jednotný formát. Tedy například *www.example.com*, *www.EXAMPLE.com* a *www.example.com/index.html* budou identifikovány jako jedna jediná stránka.

⁵Gadget je samostatná aplikace, kterou je možné jednoduše vložit na webovou stránku. Například zobrazení Google kalendáře nebo mapy.

aby jednotlivé sítě toto API podporovaly. V současné době je do projektu OpenSocial zapojeno velké množství sociálních sítí, a to včetně těch velkých jako například MySpace (kap. 3.1.8). Jediná velká síť, která do projektu zatím není zapojena, je Facebook (kap. 3.1.3).

Na první pohled vypadá tento projekt velmi slibně. Ale nevýhodou je, že většina sítí zpřístupňuje pouze informace o jednotlivých uživateli (samozřejmě pokud uživatel dá svolení). Není tedy zpřístupněn celý Sociální graf (kap. 2.4). Je však možné získat alespoň části sociálního grafu, a to díky seznamu kontaktů jednotlivých uživatelů. V současné chvíli není tedy možné jeho využití k propojení a integraci jednotlivých sociálních sítí, tak jak bychom potřebovali. Na druhou stranu by se projekt mohl hodit k získávání informací z jednotlivých sítí.

4.5 OpenId

Projekt OpenId [29] usnadňuje registraci a přihlašování do nejrůznějších webových služeb. A to nejen do sociálních sítí, i když ty jsou pro nás v této práci primární. Díky OpenId není nutné opakovaně vyplňovat registrační údaje při vytváření nových účtů. Pokud zvolená služba OpenId podporuje, stačí se identifikovat naším OpenId loginem a heslem a služba si potřebné údaje načte z našeho profilu. Spravovat a vytvářet nové OpenId účty je v dnešní době možné na velkém množství serverů.

Tento projekt neřeší vzájemné vztahy jednotlivých uživatelů. To je samozřejmě z našeho pohledu integrace velký nedostatek. Projekt však začíná být podporován více a více službami. Proto bychom OpenId mohli alespoň částečně začlenit do naší integrace. Bude však záležet na dalším zvoleném postupu.

4.6 Shrnutí

V této kapitole jsme se seznámili s několika projekty, které se pokoušejí o integraci. Žádný z nich neřeší v plné míře náš požadovaný cíl. Nicméně zmíněné projekty lze využít k řešení dílčích problémů, které v následujících kapitolách nastanou.

Kapitola 5

Co chceme integrovat

Nyní máme sociální sítě dostatečně zmapovány. Také jsme provedli průzkum projektů týkajících se integrace sociálních sítí. Na základě zjištěných informací bychom měli zvolit, jak budeme postupovat dále. Než se však pustíme do samotného návrhu integrace, je potřeba si rozmyslet, které informace vlastně budeme chtít integrovat.

V kapitole 2.3 jsme si rozebrali typickou strukturu sociální sítě. Řekli jsme si, že základem je uživatelský profil (kap. 2.3.1) a seznam kontaktů (kap. 2.3.2). Budeme tedy chtít integrovat oba dva tyto základní prvky.

Také jsme zkoumali různé vlastnosti sociálních sítí (kap. 3.2). V závěru kapitoly proto rozebereme, zda je budeme chtít také integrovat či nikoli.

5.1 Profil

Z profilů jednotlivých sítí budeme získávat profilové informace jednotlivých uživatelů. A to nejen souhrn za všechny sítě dohromady. Budeme chtít rozlišit informace podle jednotlivých sítí. Díky tomu pak budeme schopni detekovat rozdíly mezi sítěmi a upozornit na ně uživatele. Je dobré také pamatovat na fakt, že některé nesrovnalosti v profilech uživatelů mohou být záměrné. V takovém případě nebudeme uživatele upozorňovat opakovaně. S touto situací je potřeba v návrhu integrace počítat.

Integrací profilu se zabývá projekt OpenId (kap. 4.5). Jak bylo řečeno, je podporován mnoha sítěmi a další se přidávají. Na druhou stranu zatím není podporován všemi. Proto se nebudeme na projekt OpenId soustředit. Navíc při vlastním řešení budeme flexibilnější k nepředvídatelným změnám, což

u cizího projektu nelze zaručit. Pokud se však ukáže, že se OpenId rozšíří do všech sítí, není problém integraci profilu začít řešit skrze něj. Pouze bychom informace o profilu získávali jinou cestou.

5.2 Seznam kontaktů

Seznam kontaktů je jednou z nejcennějších informací, kterou jednotlivé sociální sítě mají. Kvůli tomu není možné získat sociální graf celé sítě (kap. 2.4), například pomocí API konkrétní sítě. Proto je tedy potřeba získávat dílčí sociální podgrafy jednotlivých uživatelů jinou formou. Z nich pak skládáme náš výsledný sociální graf.

Kromě celkového seznamu kontaktů jednotlivých uživatelů ze všech sítí dohromady je také vhodné vědět, jak je tomu v jednotlivých sítích. Tedy informace, že daný uživatel má někoho jiného v jedné síti ve svém seznamu kontaktů a v druhé ne nebo ho má dokonce zakázaného. Je totiž žádoucí uživatele upozorňovat právě na nesrovnalosti mezi jednotlivými sítěmi.

Nebylo by však vhodné uživatele upozorňovat neustále na to samé. Například v jedné síti má uživatel někoho ve svém seznamu kontaktů a v druhé síti ho má zakázaného. To může být způsobeno tím, že uživatel využívá jednu síť k soukromým účelům a druhou k pracovním. S touto situací musíme při návrhu počítat.

Je důležité si uvědomit, že seznam kontaktů není pouze seznamem jmen. O uživateli jsou v seznamech uvedeny i některé veřejné profilové informace. Již jsme si řekli, že máme zájem o integraci celých profilů (kap. 5.1). Může se však stát, že k nim z nějakého důvodu nebudeme mít přístup. V takovém případě bychom chtěli získat alespoň veřejné profilové informace. Zpravidla jsou ve veřejných částech profilů ta nejzásadnější data, jako je jméno a příjmení nebo email.

5.3 Vlastnosti

Dále jsme zkoumali, jaké společné vlastnosti jednotlivé sítě mají. Nejvíce rozšířenou vlastností jsou skupiny uživatelů (kap. 3.2.1). Bylo by tedy dobré se pokusit o integraci i této vlastnosti.

Velmi rozšířenou vlastností sítě je také například zasílání zpráv nebo vytváření alb fotografií. To jsou však vlastnosti, jejichž integrace by nedávala příliš velký smysl.

Z rozšiřujících vlastností se při integraci zaměříme pouze na skupiny uživatelů. Ostatními vlastnostmi se nebudeme zabývat.

5.4 Shrnutí

V této kapitole jsme rozebrali, co všechno budeme chtít v rámci sociálních sítí integrovat. Zhodnotili jsme, že budeme integrovat seznam uživatelů, včetně veřejných profilových informací, a celkové profily jednotlivých uživatelů. S tím také souvisí přihlašování do jednotlivých sítí. Musíme tedy myslet i na přístupové údaje uživatelů do jednotlivých sociálních sítí. Také jsme diskutovali rozšiřující vlastnosti sociálních sítí. Z těch jsme se rozhodli integrovat pouze skupiny uživatelů. Ostatní by nemělo příliš velký smysl integrovat.

Kapitola 6

Způsob integrace

Nyní víme, co budeme chtít integrovat. Další důležitou otázkou je, jak bude integrace zprostředkována uživateli. Dále si musíme rozmyslet, zda chceme být závislí na podpoře jednotlivých sítí (jako např. FOAF, kap. 4.1) nebo zda chceme být nezávislí.

Nejprve rozebereme, co rozumíme pojmy *závislost* a *nezávislost*. Poté zvolíme jednu z variant a svoje rozhodnutí zdůvodníme. Následně si vysvětlíme pojem *konektor*, který budeme potřebovat v dalším textu. V závěru popíšeme *Hlavní sociální síť*, která bude uživateli zprostředkovávat integraci všech ostatních sociálních sítí. Těm budeme říkat *vzdálené síť*.

6.1 Závislost vs. nezávislost

Pod pojmem závislost myslíme, že po vytvoření našeho návrhu je nutné, aby tvůrci vzdálených sítí naše řešení akceptovali a hlavně sami implementovali do své sítě. Zjistili jsme, že toto řešení není úplně nejvhodnější. Dokonce ani projekt Open Social (kap. 4.4) od tak velké společnosti jako je Google není podporován všemi, zejména tak významnou sociální sítí jako je Facebook (kap. 3.1.3). Na druhou stranu může být tento postup velmi nenáročný z pohledu naší implementace, protože zodpovědnost za implementaci do jednotlivých sítí předává autorům těchto sítí. Příkladem může být projekt FOAF (kap. 4.1).

Naopak nezávislost na vzdálených sítích si nese určité komplikace. Je nutné vytvořit postupně napojení do všech sítí. Pomocí těchto napojení lze získávat informace o kontaktech a profilech. Předchozí varianta počítala s im-

plementací stejného rozhraní v každé síti. Druhá varianta naopak počítá s rozdílnými napojeními do každé sítě. Napojení je však pouze implementační problém, což se naší práce netýká. Mnohem zajímavější pak v této variantě bude práce s informacemi získanými z jednotlivých sítí. Například pokud by informace nebyly úplné, byly rozdílné nebo pokud získáme informace různého charakteru.

Zvolíme variantu nezávislosti a tedy vlastní implementace jednotlivých napojení. Těmto napojením budeme říkat *konektory*. Přestože jejich implementace není součástí této práce, stručně si rozebereme, jak bychom je řešili.

Jedním z důvodů ke zvolení varianty nezávislosti je neexistence projektu, který by se vydal podobnou cestou. Druhým důvodem je, že ani existujícím projektům, které zvolily variantu závislosti, se nepodařilo dosáhnout cíle, který by byl uspokojivý. Je tedy lepší vyzkoušet novou variantu a v případě neúspěchu tak alespoň přinést nové informace pro budoucí projekty.

6.2 Konektory

Pro připojení do vzdálených sítí jsme si vytvořili pojem konektor. Pod jedním konektorem si představme napojení do jedné vzdálené sítě. Toto napojení nám bude zprostředkovávat požadované informace.

V kapitole 5 jsme uvedli, že chceme integrovat profily uživatelů, jejich seznam kontaktů (včetně veřejných profilových informací) a skupiny uživatelů. Všechny tyto informace nám konektor poskytne. V budoucnu by bylo vhodné konektor rozšířit o další funkce, jako například editace profilu uživatele ve vzdálené sociální síti.

Jednou z možností, jak konektor řešit, je využití stávajícího rozhraní sítě. Pokud má vzdálená síť nějaké API pro načtení profilu a seznamu kontaktů uživatele, je vytvoření konektoru velmi jednoduché. Navíc lze očekávat, že požadované informace dostaneme v jednoduše zpracovatelném formátu, obvykle formou asociativního pole nebo XML.

Velmi užitečné by mohlo být využití projektu Open Social (kap. 4.4), které nám umožní využít jedno API pro velké množství sociálních sítí.

Pokud vzdálená síť žádné rozhraní neposkytuje, jsme nuceni parsovat kód webových stránek, na kterých se nachází profilové informace a seznam kontaktů.

Je důležité poznamenat, že u všech konektorů je potřeba znát uživatelský přihlašovací údaj do vzdálené sítě. Datový model s tím tedy musí počítat. Je samozřejmé, že přihlašovací heslo bude ukládáno zašifrované.

Pod pojmem konektor si tedy můžeme představit rozhraní, které nám umožní z jedné vzdálené sítě získat profil uživatele, seznam jeho kontaktů (včetně některých jejich profilových informací) a skupiny, do kterých uživatel patří. Dohromady nám tedy všechny konektory tvoří rozhraní do všech vzdálených sítí.

6.3 Hlavní sociální síť

Nyní bychom se měli rozhodnout, jakou formou budeme integraci zprostředkovávat uživateli. Budeme se snažit vytvořit takzvanou Hlavní sociální síť, která by integrovala všechny ostatní. Naši Hlavní sociální síť budeme zkráceně říkat *Hlavní síť*.

Jak bylo řečeno v kapitole 6.2, v první fázi budeme skrz konektory získávat informace z profilů uživatelů, seznamu jejich kontaktů a některé profilové informace o kontaktech. Hlavní síť má tedy za úkol jen průběžně zjišťovat informace ze vzdálených sítí a oznamovat uživateli případné nesrovnalosti.

Tuto funkcionalitu je možné provádět bez jeho jakéhokoli pravidelného přihlašování ze strany uživatele. Postačí počáteční registrace v Hlavní síti, při které vyplní potřebné údaje včetně emailu, kam se mu mají zasílat zjištěné nesrovnalosti. Bylo by ale vhodné, aby měl uživatel možnost zobrazit své profily v jednotlivých sítích stejně tak jako seznamy svých kontaktů.

Jak jsme si definovali v kapitole 2.1, sociální síť je pro nás nějaká internetová služba, zpravidla zprostředkovaná webovým rozhraním. Proto by pro uživatele bylo pohodlné, pokud by i naše Hlavní síť byla přístupná skrz webové prostředí. Skrz něj by se registroval a zobrazoval profily a seznamy kontaktů. Také by zde měl možnost změnit svoje údaje v rámci Hlavní sítě.

6.4 Shrnutí

Zvolili jsme nezávislost na jednotlivých sociálních sítích. Budeme tedy vytvářet samostatné napojení do každé z nich. Těmto napojením budeme říkat *konektory* a ty nám umožní získat ke konkrétnímu uživateli jeho profilové informace, seznam kontaktů s některými profilovými informacemi uživatelů

v tomto seznamu a skupiny, kterých je uživatel členem. Tyto informace budeme zobrazovat v Hlavní síti, což bude webová aplikace. Nesrovnalosti v jednotlivých sítích budou uživatelům oznamovány automaticky.

Kapitola 7

Návrh integrace

V této kapitole se budeme věnovat vlastní integraci. Nejprve si určíme, jaké procesy budou probíhat v Hlavní síti (kap. 6.3), tedy jaké hlavní činnosti budou provádět uživatelé a samotná Hlavní síť. Procesy jsou částečně odvozeny od toho, co jsme se rozhodli integrovat v kapitole 5. Pokud jsou k procesu potřeba aktuální data ze vzdálených sítí, jsou načtena pomocí konektorů (kap. 6.2).

Dále identifikujeme problémy, které vznikly na základě rozebraných procesů. Každý problém popíšeme a navrhujeme jeho řešení.

V závěru kapitoly navrhujeme datový model, který by mohl být využit při vytváření Hlavní sítě. Jednotlivé části modelu budou vytvářeny na základě požadavků, vzniklých v předešlých kapitolách.

7.1 Procesy v Hlavní síti

V této kapitole identifikujeme hlavní procesy, které budou v Hlavní síti probíhat. V každé podkapitole stručně popíšeme, co daný proces obnáší.

7.1.1 Registrace uživatele

Nově registrovaný uživatel nejprve vyplní základní údaje jako jméno, příjmení, email, login a heslo. Poté je mu doporučeno, aby si nastavil systém podle svých představ. Například jestli se mu mají zasílat novinky o Hlavní síti atd. Veškerá nastavení může pak samozřejmě kdykoli v budoucnu editovat.

Velmi důležitým nastavením je, jakých vzdálených sítí je uživatel členem. Jsou mu nabídnuty na výběr ty, do kterých jsou naimplementovány konektory. Pokud je členem nějaké, která není na výběr, má možnost tuto informaci zaslat autorům Hlavní sítě. Autoři po naimplementování konektoru zveřejní zprávu o možnosti připojení do nové sítě.

Po vyplnění přístupových údajů jsou načteny jeho profily a seznamy kontaktů a jsou zobrazeny v rozhraní Hlavní sítě.

V dalším textu uvidíme (kap. 7.2.5), že někteří uživatelé mohou být založeni automaticky. Proto se může stát, že nově registrovaný uživatel je v Hlavní síti již založený. Pokud tomu tak je, je nutné mu tuto skutečnost sdělit a nezakládat nového uživatele, pouze novému uživateli umožnit editovat již založeného.

7.1.2 Synchronizace profilu

Pro každého uživatele Hlavní sítě je nutné pravidelně kontrolovat, jestli jsou jeho profily v jednotlivých vzdálených sítích shodné. Tento proces by se měl spouštět při každém přihlášení uživatele. Mnohem důležitější je však spouštět tento proces pravidelně, například každý den nebo týden, pro každého uživatele. Pokud jsou nalezeny nesrovnalosti, měl by být uživatel na tuto skutečnost upozorněn. Při kontrole po přihlášení do Hlavní sítě je mu to oznámeno přímo v rozhraní. V případě pravidelné kontroly je nejvhodnější zaslat uživateli email.

Jestli se kontrola bude provádět jen pravidelně, jen po přihlášení nebo jak po přihlášení tak pravidelně, bude možné nastavit v nastavení. Pokud bude uživatel požadovat pravidelnou kontrolu, má možnost zvolit jak často a bude mít na výběr z několika možností.

7.1.3 Synchronizace seznamu kontaktů

Podobně jako u profilů je nutné pravidelně kontrolovat nesrovnalosti v seznamech kontaktů. Proces je opět možné spouštět po každém přihlášení uživatele a také pravidelně, například každý den nebo týden. Nesrovnalosti jsou uživateli sděleny v rozhraní Hlavní sítě nebo prostřednictvím emailu. Uživatel si opět může nastavit, kdy kontrola bude probíhat a u pravidelné kontroly jak často.

7.1.4 Návrhy skupin

Stejně jako u profilu a seznamu kontaktů bude docházet ke kontrole skupin, kterých je uživatel členem. Je tu však menší rozdíl v tom, jak s informacemi naložit.

V jednotlivých vzdálených sítích vznikají rozdílné skupiny. Velmi těžko by se skupiny v různých sítích daly považovat za stejné. Ale mnohdy se dá o skupinách prohlásit, že jsou si podobné. A to dokonce i v rámci jedné vzdálené sítě. Skupiny budeme považovat za tím bližší, čím více mají stejných členů. Bylo by tedy možné uživateli nabídnout členství ve skupinách do určité hranice podobnosti. Opět by bylo možné tuto kontrolu provádět po přihlášení uživatele nebo pravidelně, s možností nastavení, kdy bude kontrola probíhat. Uživatel by si mohl nastavit i hranici podobnosti skupin.

7.1.5 Pravidelná kontrola uživatelů

Jak již bylo řečeno, někteří uživatelé mohou být založeni automaticky. Profily těchto uživatelů se budou získávat prostřednictvím konektorů uživatelů, kteří je mají ve svém seznamu kontaktů. Vzhledem k neúplnosti těchto profilů se může stát, že jeden uživatel bude automaticky založen víckrát. Přitom později může dojít k aktualizaci těchto automaticky vytvořených uživatelů. Poté by bylo jasné, že jde o toho samého uživatele.

Z těchto důvodů bude v Hlavní síti pravidelně spouštěn proces, při kterém budou kontrolovány profily všech automaticky založených uživatelů. Zkontroluje se, jestli některé automaticky založené uživatele lze ztotožnit.

7.2 Identifikace problémů a návrh řešení

V této kapitole si rozebereme, jaké problémy mohou při integraci nastat. Tyto problémy vyvstávají na základě procesů, které jsme si popsali v kapitole 7.1. V každé podkapitole rozebereme jeden konkrétní problém a navrhneme jeho řešení.

7.2.1 Ztotožnění atributů profilů ve vzdálených sítích

V rozboru procesu synchronizace profilu (kap. 7.1.2) jsme uvedli, že budeme uživatele upozorňovat na nesrovnalosti mezi profily jednotlivých vzdálených

sítí. K tomu ale potřebujeme vědět, které informace máme navzájem kontrolovat. Tedy k jakému atributu v jedné vzdálené síti přiřadit atribut ve druhé síti.

Z toho důvodu si musíme vytvořit seznam všech atributů ve všech vzdálených sítích. Každému atributu pak musíme určit, do kterých sítí patří. Pro potřeby konektorů si uložíme jeho pojmenování v jednotlivých vzdálených sítích. Také je vhodné si vést, jakého datového typu jednotlivé atributy jsou, pro lepší kontrolu jejich obsahu.

7.2.2 Uživatel chce rozdílné profily

Říkali jsme, že budeme uživatele upozorňovat na nesrovnalosti v profilech vzdálených sítí (kap. 7.1.2). Uživatel by pak měl nesrovnalosti ve vzdálených sítích opravit, jinak bychom jej upozorňovali stále dokola. Ale co když uživatel tyto nesrovnalosti chce ponechat?

Je tedy žádoucí ukládat si v Hlavní síti jednotlivé profily ze vzdálených sítí. Pak není problém případnou nesrovnalost označit za finální. Na schválené nesrovnalosti už příště nebude uživatel upozorňován. Pokud však uživatel změní dříve schválený atribut, v rámci sítě kde toto schválení provedl, je schválení odstraněno.

7.2.3 Jednoznačná identifikace uživatele

Z některých procesů (kap. 7.1) vyplývá, že je potřeba jednoznačně identifikovat uživatele. Například při pravidelné kontrole uživatelů (kap. 7.1.5). Jak takovou jednoznačnou identifikaci zajistit?

V kapitole 7.2.1 jsme došli k závěru, že si povedeme seznam atributů profilů v jednotlivých vzdálených sítích. Není potom problém některé označit za identifikátory.

Pod pojmem identifikátor si můžeme představit jakýkoli atribut, který jednoznačně identifikuje uživatele. Takovým atributem může být například email, ICQ ID, Jabber ID, kanonizovaná URL domácí stránky (kap. 4.3) atd. Tato identifikace je inspirována projektem FOAF (kap. 4.1).

7.2.4 Uživatel chce rozdílné seznamy kontaktů

Řekli jsme si, že budeme uživatele upozorňovat na nesrovnalosti v seznamech kontaktů. Ale co když tyto rozdíly uživatel požaduje? Například protože

jednu síť využívá k soukromým účelům a druhou k pracovním. Podobně jako u profilů si budeme ukládat veškerá spojení mezi uživateli v jednotlivých vzdálených sítích. Nesrovnalost zaznamenáme spojením, které označíme jako negativní. Pokud pak v dané vzdálené síti dojde ke změně a uživatelé budou v kontaktu, pouze spojení označíme jako kladné.

7.2.5 Automaticky založený uživatel

V kapitole 7.2.4 jsme si řekli, že budeme ukládat veškerá spojení mezi uživateli ve všech vzdálených sítích. Může se však stát, že uživatel načtený ze seznamu kontaktů kontrolovaného uživatele není v Hlavní síti registrovaný. V takovém případě jej jednoduše systém založí.

Je však nutné jej označit jako neregistrovaného, abychom pro něj neprováděli automatické kontroly (kap. 7.1.2 a 7.1.3). Dokonce budeme schopni vyplnit některé jeho profilové informace, i když v omezené míře (kap. 6.2).

Díky tomuto označení pak lze jednoduše provádět proces pravidelné kontroly neregistrovaných uživatelů (kap. 7.1.5). A díky alespoň částečně vyplněnému profilu lze některé uživatele jednoznačně identifikovat (kap. 7.2.3).

7.2.6 Detekce rozdílů v seznamech kontaktů

Chceme uživatele upozorňovat na nesrovnalosti v seznamech kontaktů mezi vzdálenými sítěmi (kap. 7.1.3). Jak se ale dozvíme, že kontakt kontrolovaného uživatele je členem také jiné sítě?

Jednou z možností je, že i onen kontakt je zaregistrován v Hlavní síti a víme tedy, ve kterých vzdálených sítích má účty. Další možností je, že není zaregistrovaný v Hlavní síti, ale je v ní založen automaticky (kap. 7.2.5). Potom víme, aspoň částečně, v jakých sítích má účet. V ostatních případech nelze rozhodnout.

7.2.7 Návrh nových skupin

V kapitole 7.1.4 jsme popsali proces, při kterém uživateli budeme navrhopvat členství ve skupinách. Jak se ale o nových skupinách dozvíme?

Bylo by sice možné využít vyhledávání skupin ve vzdálených sítích, ale tento proces by byl velmi náročný. Nejspíš by díky časové náročnosti nebyl v praxi vůbec použitelný. Proto si budeme skupiny, kterými jsou uživatelé

členy, ukládat v Hlavní síti. Tyto uložené skupiny pak využijeme k možným návrhům.

7.3 Návrh datového modelu

V tuhle chvíli tedy máme představu, jak bude Hlavní síť fungovat (kap. 7.1). Také jsme vyřešili problémy spojené s procesy (kap. 7.2). Nyní bychom chtěli navrhnout, jak by měl vypadat datový model Hlavní sítě.

Datový model bude tvořen na základě požadavků, vzniklých při popisu procesů a řešení problémů. V první podkapitole si rozebereme, jaké požadavky v souhrnu máme. V dalších podkapitolách na jejich základě budeme postupně tvořit datový model. Složitější celky budeme pro přehlednost zobrazovat na obrázcích. V závěru shrneme celý model včetně grafického znázornění.

Na obrázcích budou tabulky se svými atributy a vazbami. Atributy budou označeny ikonami, znázorňující primární klíč, vedlejší klíč a běžný atribut (obr. 7.1). Názvy tabulek budou v textu vyznačeny **tučně**, jejich atributy *kurzívou* a typy atributů VELKÝMI PÍSMENY.



Obrázek 7.1: Legenda k obrázkům datového modelu v kapitole 7.3.

7.3.1 Souhrn požadavků

Požadavky na datový model vznikly jak při popisu samotných procesů, tak při řešení problémů s nimi spojenými. V této kapitole si všechny sepíšeme.

Základní požadavky však vznikly již na základě konektorů (kap. 6.2). Jedná se o nutnost vedení seznamu sítí a členství uživatelů v jednotlivých

sítích včetně jejich přístupových údajů. S tím tedy souvisí nutnost vedení seznamu uživatelů. Tyto základní požadavky shrneme v jednom jediném.

Z popisu procesů v kapitole 7.1 nám vyplynul požadavek na uživatelská nastavení Hlavní sítě (kap. 7.1.1).

Z problémů v kapitole 7.2 nám vyvstaly další požadavky. Prvním je vedení seznamu atributů ve vzdálených sítích a jejich pojmenování v těchto sítích (kap. 7.2.1). Dalšími jsou ukládání profilů a seznamů kontaktů uživatelů v jednotlivých vzdálených sítích (kap. 7.2.2 a 7.2.4), možnost odlišení neregistrovaných uživatelů (kap. 7.2.5) a jednoznačná identifikace uživatele (kap. 7.2.3). Posledním požadavkem je ukládání skupin v jednotlivých sítích (kap. 7.2.7).

Nyní máme sepsány požadavky na datový model. V dalších kapitolách budeme postupně tvořit výsledný návrh datového modelu na jejich základě.

7.3.2 Uživatelé, síť a členství

V Hlavní síti si musíme vést seznam jejich členů. K tomu bude sloužit tabulka **user**. Budou v ní uloženy základní informace o uživateli, jako jsou například jejich přihlašovací údaje do Hlavní sítě. Atribut *email* slouží k uložení emailové adresy, na kterou budeme uživateli zasílat upozornění a novinky Hlavní sítě.

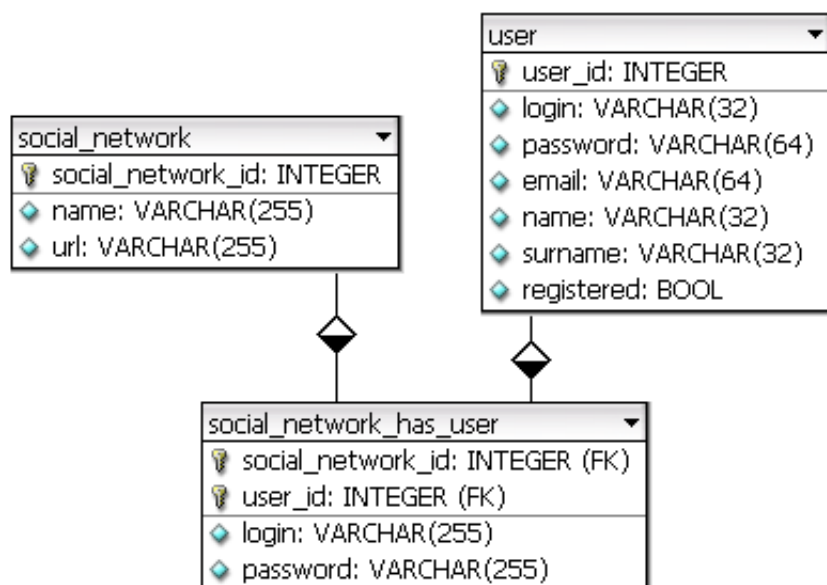
V tabulce **social_network** bude uložen seznam jednotlivých vzdálených sítí. V tuto chvíli nám postačí jméno sítě pro zobrazení v Hlavní síti a její URL pro potřeby konektorů.

Členství uživatelů v jednotlivých vzdálených sítích, včetně jejich přihlašovacích údajů, je vyřešeno tabulkou **social_network_has_user**.

Hesla do vzdálených sítí i do Hlavní sítě by měla být z bezpečnostních důvodů ukládána šifrovaně. Tabulky jsou znázorněny na obrázku 7.2. Atribut *registered* v tabulce **user** bude odůvodněn v kapitole 7.3.3.

7.3.3 Neregistrovaní uživatelé

Jak víme, potřebujeme rozlišit uživatele, kteří se registrovali do Hlavní sítě, od těch, kteří byli založeni automaticky. K tomu bude sloužit atribut *registered* typu **BOOL** v tabulce **user**. Registrovaní uživatelé jej budou mít nastaven na **TRUE**, neregistrovaní na **FALSE**.



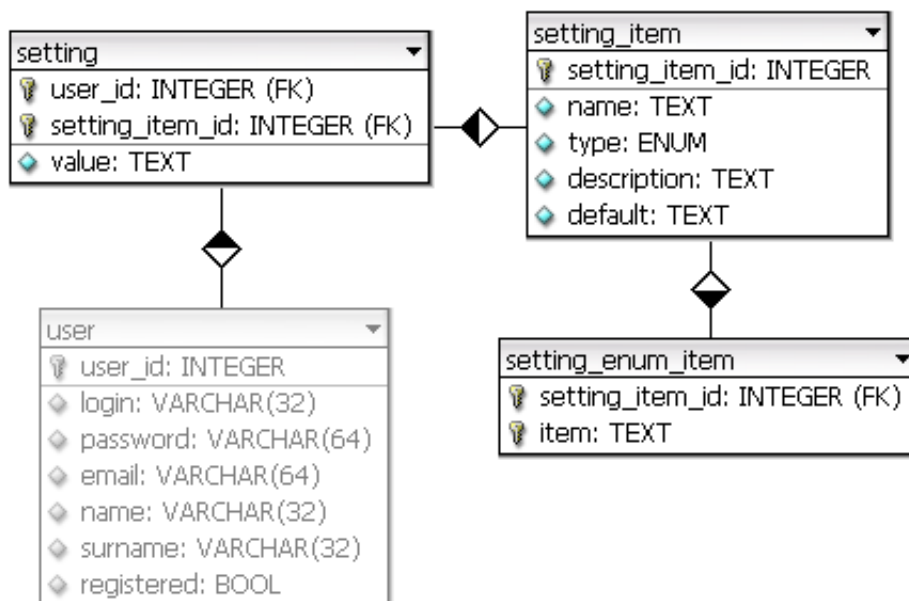
Obrázek 7.2: Část modelu řešící uložení základních informací o uživateli a vzdálených sítích. Atribut *registered* v tabulce **user** určuje, zda je uživatel zaregistrovaný v Hlavní síti nebo byl vytvořen automaticky. V tabulce **social_network_has_user** jsou uloženy přihlašovací údaje uživatelů do jednotlivých vzdálených sítí. Atributy *password* v tabulkách **user** a **social_network_has_user** by měly být ukládány šifrovaně.

7.3.4 Nastavení

Seznam možných nastavení bude uložen v tabulce **setting_item**. Atributy *name* a *description*, ve kterých je uloženo jméno a popis nastavení, jsou pro zobrazení v Hlavní síti. Typ dané položky je uveden v atributu *type* (např. TEXT, INTEGER, ENUM, DATE atd.). V tabulce **setting_enum_item** jsou uloženy jednotlivé hodnoty pro výčtový typ ENUM.

Nastavení jednotlivých uživatelů budeme ukládat v tabulce **setting**. Hodnota daného nastavení (z tabulky **setting_item**) bude uložena v atributu *value* typu TEXT. Dokud uživatel nenastaví hodnotu sám, použije se pro nastavení výchozí hodnota. Ta je pro každý typ nastavení uložena v atributu *default* (tabulka **setting_item**).

Tabulky jsou znázorněny na obrázku 7.3.



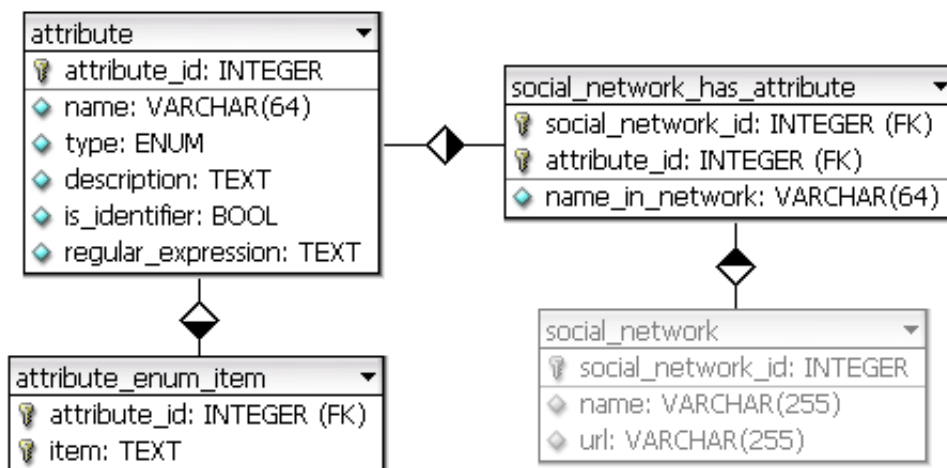
Obrázek 7.3: Tabulky modelu, sloužící k uložení uživatelských nastavení. V tabulce **setting_item** je seznam všech možných nastavení. V tabulce **setting** jsou nastavení pro jednotlivé uživatele. V tabulce **setting_enum_item** jsou uloženy možné hodnoty nastavení, které mají nastavený typ na ENUM (atribut *type* v tabulce **setting_item**). Je zde také světle zobrazena již zmíněná tabulka **user**, aby bylo možné zobrazit všechny vzájemné vazby.

7.3.5 Atributy profilů

Seznam všech atributů ve všech vzdálených sítích bude uložen v tabulce **attribute**. Pro zobrazení v Hlavní síti jsou zde atributy *name* a *description*, ve kterých je uloženo jméno a popis atributu. Atribut *type* určuje, jakého typu daný atribut je (např. TEXT, INTEGER, ENUM, DATE atd.). V tabulce **attribute_enum_item** jsou uloženy jednotlivé hodnoty pro výčtový typ ENUM.

Do tabulky **social_network_has_attribute** budeme zaznamenávat, které atributy se ve kterých sítích vyskytují. Jméno atributu ve vzdálené síti je uloženo v této tabulce v atributu *name_in_network*.

Tabulky jsou znázorněny na obrázku 7.4. Atributy *is_identifier* a *regular_expression* v tabulce **social_network_has_attribute** budou vysvětleny až v kapitole 7.3.7.



Obrázek 7.4: Tato část modelu slouží k uložení profilových atributů v jednotlivých vzdálených sítích. V tabulce **attribute** je seznam všech atributů ve všech vzdálených sítích. Atribut *is_identifier* typu BOOL slouží k označení identifikátorů, což jsou významné profilové atributy. Výskyt atributů v jednotlivých vzdálených sítích zaznamenává tabulka **social_network_has_attribute**. V tabulce **attribute_enum_item** jsou uloženy možné hodnoty atributů, které mají nastavený typ na ENUM (atribut *type* v tabulce **attribute**). Je zde také světle zobrazena již zmíněná tabulka **social_network**, aby bylo možné zobrazit všechny vzájemné vazby.

7.3.6 Profily

Profilové informace jednotlivých uživatelů ve vzdálených sítích budou uloženy v tabulce **profile**. Každá hodnota bude asociována s uživatelem, vzdálenou sítí a atributem profilu. Hodnota bude uložena v atributu *value*.

Odlišné hodnoty profilu, schválené uživatelem, budou označeny. K tomu bude sloužit atribut *final* typu BOOL. Společné hodnoty ve vzdálených sítích budou uloženy bez vazby na vzdálenou síť. Pro odlišnou hodnotu v konkrétní

vzdálené síti bude vytvořen záznam s vazbou na požadovanou vzdálenou síť a atribut *final* bude nastaven na TRUE.

7.3.7 Identifikátory

Některé profilové atributy budeme chtít označit jako identifikátory. K tomu bude sloužit atribut *is_identifier* typu BOOL v tabulce **attribute**.

Identifikátory jsou velmi důležité atributy profilu uživatele. Slouží k jeho jednoznačné identifikaci. Hodnoty atributů jsou kontrolovány pomocí datového typu (atribut *type*), ale to nám u identifikátorů nestačí. Proto je v tabulce **attribute** atribut *regular_expression*. V něm je uložen regulární výraz, kterým každou ukládanou hodnotu zkontrolujeme. Příkladem může být emailová adresa, jejíž hodnotu pouhým datovým typem nekontrolujeme.

7.3.8 Seznamy kontaktů

Vzájemná spojení mezi uživateli ve vzdálených sítích budeme zaznamenávat do tabulky **connection**. Pro každé dva uživatele je vytvořeno maximálně jedno spojení. Ve kterých vzdálených sítích je konkrétní spojení realizováno, je uloženo v tabulce **social_network_has_connection**.

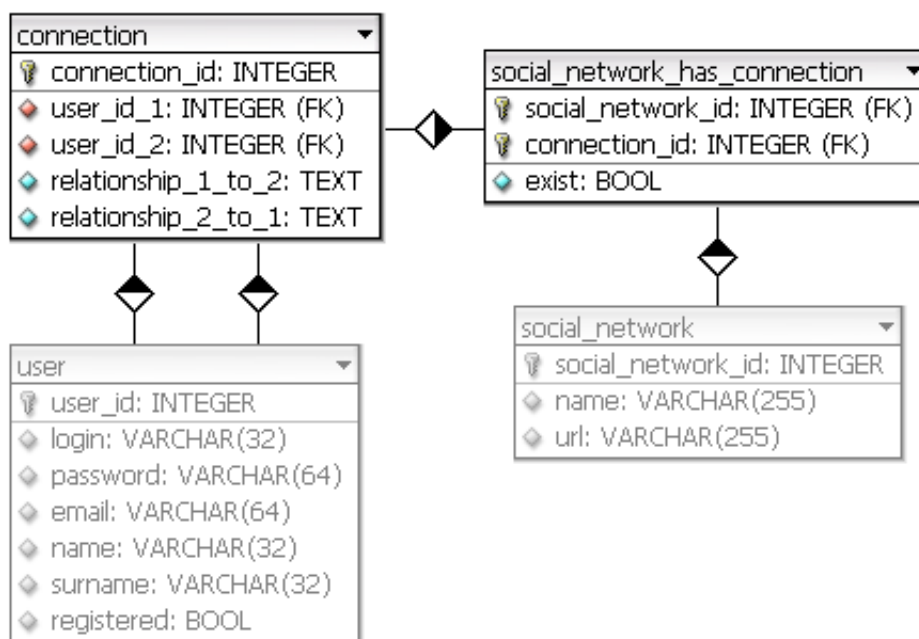
Uživatel si může explicitně vyžádat rozdílnost v kontaktech některé vzdálené síť. K tomu je určen atribut *exist*. Explicitně záporné spojení, tedy že uživatelé nejsou v dané vzdálené síti v kontaktu, bude mít tento atribut nastaveno na FALSE.

V tabulce **connection** jsou ještě dva atributy *relationship_1_to_2* a *relationship_2_to_1*. Tyto atributy slouží k zaznamenání, jaké vztahy mezi uživateli jsou a to oběma směry. Informace z těchto atributů mohou být v budoucnu užitečné pro analýzu uživatelských vztahů a preferencí. Tyto atributy byly inspirovány projektem XFN (kap. 4.2).

Tabulky jsou znázorněny na obrázku 7.5.

7.3.9 Skupiny

Skupiny uživatelů budeme ukládat do tabulky **group**. Každá skupina má určeno, do které vzdálené sítě patří. Skrz konektor je ze vzdálené sítě převzat název a popis skupiny (atributy *name* a *description*). Atribut *is_public* typu BOOL určuje, jestli je skupina veřejná. Tuto informaci získáme také skrz konektor.

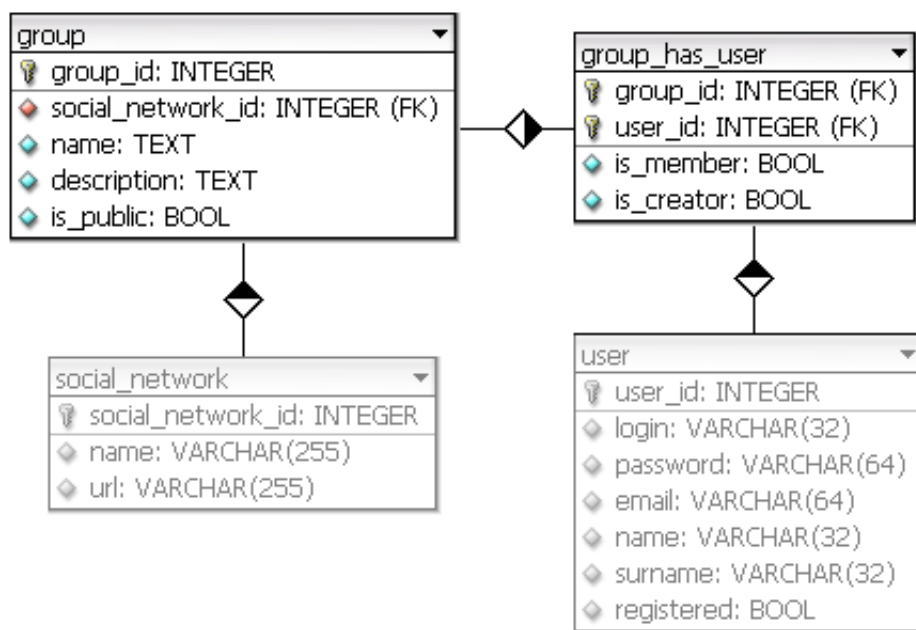


Obrázek 7.5: Část modelu, která řeší vztahy mezi uživateli v jednotlivých vzdálených sítích. V tabulce **connection** jsou kromě referencí na dva uživatele, kteří jsou v nějakém vztahu, uloženy i jejich vzájemné vztahy v atributech *relationship_1_to_2* a *relationship_2_to_1*. Tabulka **social_network_has_connection** zaznamenává existenci vztahu mezi dvěma uživateli (definovanou v tabulce **connection**) v jednotlivých sítích. Díky atributu *exist* je také možné explicitně zaznamenat neexistenci vztahu ve zvolené síti. Jsou zde také světle zobrazeny již zmíněné tabulky **user** a **social_network**, aby bylo možné zobrazit všechny vzájemné vazby.

V tabulce **group_has_user** je uloženo členství uživatelů v jednotlivých skupinách. Atribut *is_creator* určuje, zda je příslušný uživatel zakladatelem této skupiny.

Pokud bychom uživateli doporučili členství ve skupině, jejímž členem nechce být, vytvoříme záznam v tabulce **group_has_user** a nastavíme atribut *is_member* na **FALSE**.

Tabulky jsou znázorněny na obrázku 7.6.



Obrázek 7.6: Část modelu, která řeší skupiny uživatelů v jednotlivých vzdálených sítích. V tabulce **group** jsou uloženy jednotlivé skupiny. Členství ve skupinách je zaznamenáno v tabulce **group_has_user**. Pokud chceme explicitně zaznamenat, že uživatel není členem skupiny, nastavíme atribut *is_member* na FALSE. Jsou zde také světle zobrazeny již zmíněné tabulky **user** a **social_network**, aby bylo možné zobrazit všechny vzájemné vazby.

7.3.10 Celý model

Postupně jsme vyřešili veškeré požadavky, vyjmenované na začátku této kapitoly. Model obsahuje základní část pro správu uživatelů a vzdálených sociálních sítí. Dále část obstarávající uživatelská nastavení a správu profilových atributů ve vzdálených sítích. Jsou zde také části ukládající profily uživatelů, jejich seznamy kontaktů a členství ve skupinách. Jsme schopni jednoznačně identifikovat uživatele a rozlišit registrované uživatele od neregistrovaných.

Na obrázku 7.7 je zobrazen celý model se všemi vazbami a tabulkami.

stavení systému podle svých představ. Procesy *synchronizace profilu* a *synchronizace seznamu kontaktů* obnášejí kontrolu nesrovnalostí ve vzdálených sítích. Při procesu *návrhy skupin* je uživatel upozorněn na skupiny blízké těm, ve kterých už je členem. Posledním procesem je *pravidelná kontrola uživatelů*, při kterém se hledají a odstraňují duplicity mezi automaticky založenými uživateli.

V druhé části kapitoly jsme identifikovali problémy plynoucí z uvedených procesů. U všech jsme však předložili jejich řešení.

V závěru kapitoly jsme navrhli datový model, který by bylo možné použít při vytváření Hlavní sítě. Jednotlivé části modelu jsme vytvářeli na základě požadavků, vzniklých v předchozích kapitolách.

Kapitola 8

Závěr

V této práci jsme měli za úkol navrhnout způsob integrace informací z jednotlivých sociálních sítí. Přínos by z ní měli jak uživatelé, tak provozovatelé sociálních sítí. Z pohledu uživatelů je základním motivem přehlednost dat a jednoduchost práce s těmito daty. Z pohledu provozovatelů je zřejmý ekonomický přínos díky datům, která získají o uživatelích.

Na úvod jsme se seznámili se základními pojmy, abychom se lépe orientovali v problematice. Nadefinovali jsme si, že pod sociální sítí budeme chápat webovou službu, pomocí které je možné vytvářet vazby a sdílet informace mezi jednotlivými uživateli. Řekli jsme si, jaké jsou typické rysy sociální sítě. Prvním z nich je možnost editace uživatelského *profilu*. Druhým je možnost vytváření *seznamu kontaktů*, tedy vytváření vazeb mezi jednotlivými uživateli. Vazby mezi uživateli tvoří graf, kterému se v oblasti sociálních sítí říká *sociální graf*.

V další části jsme provedli rozbor existujících sociálních sítí. Nejprve jsme si vyjmenovali některé existující sítě a stručně jsme si je představili. Ve výčtu jsme se zaměřili na nejznámější sociální sítě a také na některé zajímavé svým zaměřením. Abychom měli o sítích větší přehled, prozkoumali jsme, jaké mají rozšiřující vlastnosti. Nejrozšířenějšími vlastnostmi byly *skupiny uživatelů* a *psaní zpráv mezi uživateli*.

Dále jsme rozebrali dvě zásadní kritéria, podle kterých lze dělit sociální sítě. Prvním je *zaměření*. To nám sociální sítě dělí na *obecné* a *specializované*. Hlavním cílem obecných sociálních sítí je udržování kontaktů mezi přáteli, kolegy nebo obchodními partnery. Specializované se soustředí na určitou oblast, například hudbu, knihy nebo kariéru.

Téma integrace sociálních sítí není úplně nové. Proto existují projekty, které se tímto tématem zabývají. Některé z těchto projektů jsme prozkoumali a popsali, abychom se z nich poučili a případně vyvarovali některých chyb. Při zkoumání projektů jsme nenašli žádný, který by v plné míře splňoval naše požadavky. Přesto bylo jejich zkoumání přínosné, jak se ukázalo v dalším textu. Z některých jsme převzali zajímavé myšlenky a některé by dokonce bylo možné využít při implementaci.

Po seznámení s pojmy, rozboru existujících sítí a existujících projektů, zabývajících se podobnou problematikou, jsme si stanovili, co budeme chtít integrovat. Základními prvky sociálních sítí jsou uživatelský profil a seznam kontaktů. V rozboru vlastností sociálních sítí jsme zjistili, že nejrozšířenější vlastností jsou skupiny uživatelů. Proto jsme se rozhodli tyto tři prvky integrovat. Ostatní vlastnosti buď nemá příliš velký význam integrovat, nebo nejsou natolik rozšířené, abychom se jimi zabývali.

Dále jsme si stanovili, jak budeme k integraci přistupovat. Diskutovali jsme dva přístupy, závislost a nezávislost na jednotlivých sociálních sítích. Rozhodli jsme se pro nezávislost. Učinili jsme tak na základě zkušeností, které jsme načerpali při zkoumání existujících projektů. Vzhledem k této volbě jsme si zavedli pojem *konektor*. Jedná se o napojení do jednotlivých sociálních sítí, které nám zprostředkovává informace o uživateli.

Také jsme se rozhodli, jakým způsobem budeme integraci zprostředkovávat uživatelům. Bude nám k tomu sloužit webová aplikace, kterou nazveme *Hlavní síť*. V této souvislosti jsme si označili ostatní sociální sítě jako *vzdálené sítě*. Hlavní síť bude sloužit k zobrazení veškerých informací ze vzdálených sítí a oznamování případných nesrovnalostí v profilech a seznamech kontaktů uživatelů. Také bude navrhopvat členství ve skupinách podobných těm, kterými jsou již uživatelé členy.

V závěrečné kapitole jsme přednesli samotný návrh integrace. Nejprve jsme si stanovili, jaké procesy budou probíhat v Hlavní síti. Na základě těchto procesů nám vyvstaly určité problémy. Postupně jsme si je popsali a navrhli jejich řešení.

Následně jsme navrhli datový model Hlavní sítě. Tvořili jsme jej na základě požadavků, které vznikly v souvislosti s procesy a problémy v předchozím textu. Většina požadavků má v zásadě jedno společné. Snaha veškeré informace ukládat do Hlavní sítě, aby bylo možné zaznamenat rozdíly schválené uživatelem.

Samotná implementace nebyla cílem práce. Jejím cílem byl návrh integrace a tento cíl jsme splnili. Navrhli jsme model Hlavní sítě a identifikovali

hlavní procesy, které v ní budou nastávat. Informace ze vzdálených sítí budou získávány prostřednictvím konektorů. Jejich implementace je závislá na možnostech jednotlivých vzdálených sítí. V práci jsme však zmínili možná řešení jejich implementace, včetně návrhu částečně obecného řešení prostřednictvím existujícího projektu.

Kapitola 9

Budoucí práce

V budoucnu by bylo možné na tuto práci navázat jak implementačně tak teoreticky. Proto nastíníme, jak by se dalo pokračovat.

Implementační navázání je zřejmé. Bude nutné vytvořit nejen samotnou Hlavní síť, ale také jednotlivé konektory do vzdálených sítí. Mohlo by být zajímavé pokusit se o zobrazení sociálního grafu v rámci Hlavní sítě s vyznačením oblastí, které spadají pod jednotlivé vzdálené sítě.

Pokračování v teoretické části by mohlo rozvinout navrhování podobných skupin. V textu jsme uváděli, že podobné skupiny budeme rozpoznávat na základě jejich členů. Netriviálním problémem by bylo navržení algoritmu, který by označil skupiny za podobné i na základě jiných informací. Například jejich názvu, popisu nebo jiných atributů. V této souvislosti by nejspíš bylo vhodné rozšířit i datový model.

Další teoretickou oblastí, kde by bylo možné navázat, by byla identifikace uživatele. V práci jsme k tomuto účelu využívali identifikátory. Bylo by zajímavé vytvořit algoritmy, které identifikují uživatele i bez identifikátorů. Nejspíš by bylo možné tuto identifikaci provádět pouze s určitou pravděpodobností.

Velmi praktické rozšíření, které by bylo možné pojmout jak teoreticky tak implementačně, jsou statistiky. Jak bylo řečeno, pro obchodníky jsou důležitá data uživatelů. Aby je bylo možné využít, je potřeba jejich statistické zpracování. Z teoretického hlediska by bylo potřeba navrhnout zpracování dat uložených v Hlavní síti. Z implementačního hlediska by bylo nutné navrhnout jak data smysluplně zobrazovat, případně filtrovat.

Toto byly pouze určité návrhy, jak v práci pokračovat. Zajisté by se našla celá řada dalších nápadů, jak teoretických, tak implementačních.

Literatura

- [1] *ARPANET - Wikipedia, the free encyclopedia* [online], cit. 9.11.2008.
Dostupný z WWW: <http://en.wikipedia.org/wiki/ARPANET>
- [2] *Usenet - Wikipedia, the free encyclopedia* [online], cit. 9.11.2008.
Dostupný z WWW: <http://en.wikipedia.org/wiki/Usenet>
- [3] *SixDegrees.com - Wikipedia, the free encyclopedia* [online], cit. 9.11.2008.
Dostupný z WWW: <http://en.wikipedia.org/wiki/SixDegrees.com>
- [4] *Friendster* [online], cit. 9.11.2008. Dostupný z WWW:
<http://www.friendster.com/>
- [5] *MySpace* [online], cit. 9.11.2008. Dostupný z WWW:
<http://www.myspace.com/>
- [6] *Social network service - Wikipedia, the free encyclopedia* [online],
cit. 9.11.2008. Dostupný z WWW:
http://en.wikipedia.org/wiki/Social_network_service
- [7] *Categorized List of Social Media and Social Networking Websites - Traffikd* [online], cit. 9.11.2008. Dostupný z WWW:
<http://traffikd.com/social-media-websites/>
- [8] *List of social networking websites - Wikipedia, the free encyclopedia* [online], cit. 9.11.2008. Dostupný z WWW:
http://en.wikipedia.org/wiki/List_of_social_networking_websites
- [9] *CSFD: Česko-Slovenská filmová databáze* [online], cit. 20.2.2009.
Dostupný z WWW: <http://www.csfd.cz/>
- [10] *aNobii: Create, Share, and Explore booklists* [online], cit. 5.4.2009.
Dostupný z WWW: <http://www.anobii.com/>

- [11] *CouchSurfing - Participate in Creating a Better World, One Couch At A Time* [online], cit. 10.4.2009. Dostupný z WWW: <http://www.couchsurfing.com/>
- [12] *Facebook* [online], cit. 6.4.2009. Dostupný z WWW: <http://www.facebook.com/>
- [13] *ASEAN, Association of South East Asian Nations* [online], cit. 10.4.2009. Dostupný z WWW: <http://www.asean.org/>
- [14] *Last.fm - Listen to internet radio and the largest music catalogue online* [online], cit. 12.4.2009. Dostupný z WWW: <http://www.last.fm/>
- [15] *Lide.cz* [online], cit. 5.4.2009. Dostupný z WWW: <http://www.lide.cz/>
- [16] *LinkedIn* [online], cit. 12.4.2009. Dostupný z WWW: <http://www.linkedin.com/>
- [17] *PatientsLikeMe : Patients Helping Patients Live Better Every Day* [online], cit. 13.4.2009. Dostupný z WWW: <http://www.patientslikeme.com/>
- [18] *Scientific Network - ResearchGATE* [online], cit. 12.4.2009. Dostupný z WWW: <http://www.researchgate.net/>
- [19] *SoberCircle Recovery Community* [online], cit. 12.5.2009. Dostupný z WWW: <http://www.sobercircle.com/>
- [20] *The Friend of a Friend (FOAF) project - FOAF project* [online], cit. 1.3.2009. Dostupný z WWW: <http://www.foaf-project.org/>
- [21] *RDF/XML Syntax Specification (Revised)* [online], cit. 1.3.2009. Dostupný z WWW: <http://www.w3.org/TR/rdf-syntax-grammar/>
- [22] *OWL Web Ontology Language* [online], cit. 1.3.2009. Dostupný z WWW: <http://www.w3.org/TR/owl-ref/>
- [23] *W3C Semantic Web Activity* [online], cit. 1.3.2009. Dostupný z WWW: <http://www.w3.org/2001/sw/>
- [24] *XFN - XHTML Friends Network* [online], cit. 18.4.2009. Dostupný z WWW: <http://gmpg.org/xfn/>

- [25] *Mikroformát - Wikipedie, otevřená encyklopedie* [online], cit. 11.5.2009.
Dostupný z WWW: <http://cs.wikipedia.org/wiki/Mikroformáty>
- [26] *Social Graph API - Google Code* [online], cit. 1.5.2009.
Dostupný z WWW: <http://code.google.com/apis/socialgraph/>
- [27] *google-sgnodemapper - Project Hosting on Google Code* [online],
cit. 1.5.2009. Dostupný z WWW:
<http://code.google.com/p/google-sgnodemapper/>
- [28] *OpenSocial - Google Code* [online], cit. 10.5.2009. Dostupný z WWW:
<http://code.google.com/intl/cs/apis/opensocial/>
- [29] *OpenID Foundation website* [online], cit. 18.4.2009. Dostupný z WWW:
<http://www.openid.net/>